

# RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following CRF diskette:

Application Serial Number 09/029,042A  
Art Unit / Team No. 1646  
Date Processed by STIC 10/28/99

THE ATTACHED PRINTOUT EXPLAINS THE ERRORS DETECTED.

PLEASE BE SURE TO FORWARD THIS INFORMATION TO THE APPLICANTS BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANTS ALONG WITH A NOTICE TO COMPLY or,
- 2) CALLING APPLICANTS AND FAXING THEM A COPY OF THE PRINTOUT WITH A NOTICE TO COMPLY

THIS WILL INSURE THAT THE NEXT SUBMISSION RECEIVED FROM THEM WILL BE ERROR FREE.

IF YOU HAVE ANY FURTHER QUESTIONS, PLEASE CALL:

MARK SPENCER 703-308-4212

# Raw Sequence Listing Error Summary

<u>ERROR DETECTED</u>	<u>SUGGESTED CORRECTION</u>	<u>SERIAL NUMBER:</u> <u>09/029,042A</u>
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 <input type="checkbox"/> Wrapped Nucleics	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3, as this will prevent "wrapping".	
2 <input type="checkbox"/> Wrapped Aminos	The amino acid number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3, as this will prevent "wrapping".	
3 <input type="checkbox"/> Incorrect Line Length	The rules require that a line not exceed 72 characters in length. This includes spaces.	
4 <input type="checkbox"/> Misaligned Amino Acid Numbering	The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs between the numbering. It is recommended to delete any tabs and use spacing between the numbers.	
5 <input type="checkbox"/> Non-ASCII	This file was not saved in ASCII (DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text so that it can be processed.	
6 <input type="checkbox"/> Variable Length	Sequence(s) _____ contain n's or Xaa's which represented more than one residue. As per the rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the (ix) feature section that some may be missing.	
7 <input type="checkbox"/> PatentIn ver. 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequence(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence.	
8 <input type="checkbox"/> Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS") (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: <b>This sequence is Intentionally skipped</b>  Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).	
9 <input type="checkbox"/> Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please use the following format for each skipped sequence. <210> sequence id number <400> sequence id number 000	
10 <input type="checkbox"/> Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Use of <220> to <223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
11 <input type="checkbox"/> Use of <213>Organism (NEW RULES)	Sequence(s) _____ are missing this mandatory field or its response.	
12 <input type="checkbox"/> Use of <220>Feature (NEW RULES)	Sequence(s) _____ are missing the <220>Feature and associated headings. Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown" Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)	
13 <input type="checkbox"/> PatentIn ver. 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other means to copy file to floppy disk.	

D Fitzgerald

1646

PAGE: 1

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/029,042A

DATE: 10/28/1999  
TIME: 14:00:56

Input Set: I029042A.RAW

This Raw Listing contains the General Information Section and those Sequences containing ERRORS.

Does Not Comply  
Corrected Diskette Needed

1 <110> Kim, Sun-Young; Kim, Kee-Won; Kim, Tae-Han; Hwang, Jeong-Ho; Kim, Seon-Hee; L  
2 <120> Heterologous Protein Production System using Avian Cells  
3 <130>  
4 <140> US/09/029,042A  
5 <141> 1998-05-15  
6 <150> PCT/KR96/00145  
7 <151> 1996-08-23  
8 <160> 11

Please consult  
new Sequence Rules  
and Sample Sequence  
Listings (attached in back)  
for valid format.

all text  
must be  
visible -  
See item 3  
on Error  
Summary Sheet

ERRORED SEQUENCES FOLLOW

9 <210> 1  
10 <211> 1584  
11 <212> DNA  
12 <213> erythropoietin → 1  
E--> 13 <400> ATGGGGTGC ACGAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCCCTGCT  
E--> 14 GTCGCTCCCT CTGGGCTCC CAGTCCTGGG CGCCCCACCA CGCCTCATCT  
W--> 15  
E--> 16 GTGACAGCCG AGTCCTGGAG AGGTACCTCT TGGAGGCCAA GGAGGCCAG  
W--> 17 AATATCACGG TGAGACCCCT TCCCCAGCAC ATTCCACAGA ACTCACGCTC  
E--> 18 dashes are INVALID  
W--> 19 AGGGCTTCAG GG-AACTCCT CCCAG-ATCC AGGAACCTGG CACTTGGTTT  
E--> 20  
W--> 21 GGGGTGGAGT TGGGAAGCTA GACACTGCC CCCTACATAA GAATAAGTCT  
E--> 22  
W--> 23 GGTGGCCCCA AACCATAACCT GGAAACTAGG CAAGGAGCAA AGCCAGCAGA  
E--> 24  
W--> 25 TCCTACGGCC TGTGGGCCAG GGCCAG(AGC CTTCAGGGAC CCTTGACTCC  
E--> 26  
W--> 27 CCGGGCTGTG TGCATTCAG ACGGGCTGTG CTGAACACTG CAGCTTGAAT  
E--> 28  
W--> 29 GAGAATATCA CTGTCCCAGA CACCAAAGTT AATTCTATG CCTGGAAGAG  
E--> 30  
W--> 31 GATGGAGGTG AGTTCCCTTT TTTTTTTTT TCCTTTCTTT TGGAGAATCT  
E--> 32  
W--> 33 CATTTGCGAG CCTGATTTG GATGAAAGGG AGAATGATCG GGGGAAAGGT  
E--> 34  
W--> 35 AAAATGGAGC AGCAGAGATG AGGCTGCCTG GGCGCAGAGG CTCACGTCTA  
E--> 36  
W--> 37 TAATCCCAGG CTGAGATGGC CGAGATGGGA GAATTGCTTG AGCCCTGGAG  
E--> 38  
W--> 39

insert sequence number NEXT TO <400>  
move bases down - DO NOT USE <400> line

100  
50

Per 1822  
of sequence  
Rules, insert  
the cumulative  
base total at  
end of each  
line.

DO NOT use  
upper-case  
letters for  
nucleic acids;  
under new  
Sequence Rules,  
use lower-case  
letters for bases.

PAGE: 2

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/029,042ADATE: 10/28/1999  
TIME: 14:00:56

Input Set: I029042A.RAW

E--> 40 GTTCAGACCA ACCTAGGCAG CATA GTGAGA TCCCCATCT CTACAAACAT  
 W--> 41 TTAAAAAAAT TAGTCAGGTG AAGTGGTGCA TGGTGGTAGT CCCAGATATT  
 E--> 42 TGGAAGGCTG AGGCAGGGAGG ATCGCTTGAG CCCAGGAATT TGAGGCTGCA  
 W--> 43 GTGAGCTGTG ATCACACCAC TGCACTCCAG CCTCAGTGAC AGAGTGAGGC  
 E--> 44 CCTGTCTCAA AAAAGAAAAG AAAAAAGAAA ATAATGAGG GCTGTATGGA  
 W--> 45 ATACATTCA TATTCA CTCACTCACT CACTCATTCA TTCATTCAATT  
 E--> 46 CATTCAACAA GTCTTATTGC ATACCTCTG TTTGCTCAGC TTGGTGCTTG  
 W--> 47 GGGCTGCTGA GGGGCAGGAG GGAGAGGGTG ACATGGGTCA GCTGACTCCC  
 E--> 48 AGAGTCCACT CCCTGTAGGT CGGGCAGCAG GCCGTAGAAG TCTGGCAGGG  
 W--> 49 CCTGGCCCTG CTGTCGGAAG CTGTCCTGCG GGGCCAGGCC CTGTTGGTCA  
 E--> 50 ACTCTTCCA GCCGTGGGAG CCCCTGCAGC TGCATGTGGA TAAAGCCGTC  
 W--> 51 AGTGGCCTTC GCAGCCTCAC CACTCTGCTT CGGGCTCTGG GAGCCCAGGT  
 E--> 52 GAGTAGGAGC GGACACTTC GCTTGCCTT TCTGTAAGAA GGGGAGAAGG  
 W--> 53 GTCTTGCTAA GGAGTACAGG AACTGTCCGT ATTCCCTCCC TTTCTGTGGC  
 E--> 54 ACTGCAGCGA CCTCCTGTT TCTCCTTGGC AGAAGGAAGC CATCTCCCCT  
 W--> 55 CCAGATGCAGG CCTCAGCTGC TCCACTCCGA ACAATCACTG CTGACACTTT  
 E--> 56 CCGCAAACTC TTCCGAGTCT ACTCCAATTT CCTCCGGGAA AAGCTGAAGC  
 W--> 57 TGTACACAGG GGAGGCCTGC AGGACAGGGG ACAGATGA  
 E--> 58  
 W--> 59  
 E--> 60  
 W--> 61  
 E--> 62  
 W--> 63  
 E--> 64  
 W--> 65  
 E--> 66  
 W--> 67  
 E--> 68  
 W--> 69  
 E--> 70  
 W--> 71  
 E--> 72  
 W--> 73  
 E--> 74  
 W--> 75

*present*  
*Cumulative*  
*base dots*

76 <210> 2  
 77 <211> 1582  
 78 <212> DNA  
 79 <213> erythropoietin 2  
 E--> 80 (400) ATGGGGGTGC ACGAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCCCTGCT  
 E--> 81 GTCGCTCCCT CTGGGCCTCC CAGTCCTGGG CGCCCCACCA CGCCTCATCT  
 W--> 82  
 E--> 83 GTGACAGCCG AGTCCTGGAG AGGTACCTCT GGAGGCCAAG GAGGCCGAG  
 W--> 84  
 E--> 85 AATATCACGG TGAGACCCCT TCCCCAGCAC ATTCCACAGA ACTCACGCTC  
 W--> 86 AGGGCTTCAG GG(ACTCCT CCCAG ATCC AGGAACCTGG CACTTGGTTT  
 E--> 87  
 W--> 88

*same errors*

PAGE: 3

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/029,042ADATE: 10/28/1999  
TIME: 14:00:56

Input Set: I029042A.RAW

E--> 89 GGGGTGGAGT TGGGAAGCTA GACACTGCC CCCTACATAA GAATAAGTCT  
W--> 90  
E--> 91 GGTGGCCCCA AACCATAACCT GGAAACTAGG CAAGGAGCAA AGCCAGCAGA  
W--> 92 TCCTACGCC TGTGCAGGCCAGGGAGC CTTCAGGGAC CCTTGACTCC  
E--> 93  
W--> 94 CCGGGCTGTG TGCATTTCAAG ACAGGGCTGTG CTGAACACTG CAGCTTGAAT  
E--> 95  
W--> 96 GAGAATATCA CTGTCCCAGA CACCAAAGTT AATTCTATG CCTGGAAGAG  
E--> 97  
W--> 98 GATGGAGGTG AGTTCCCTTT TTTTTTTTT TCCTTTCTTT TGGAGAATCT  
E--> 99 CATTTGCGAG CCTGATTTTG GATGAAAGGG AGAATGATCG AGGGAAAGGT  
W--> 100 AAAATGGAGC AGCAGAGATG AGGCTGCCTG GGCGCAGAGG CTCACGTCTA  
E--> 101 TAATCCCAGG CTGAGATGGC CGAGATGGGA GAATTGCTTG AGCCCTGGAG  
W--> 102 GTTCAGACCA ACCTAGGCAG CATACTGAGA TCCCCCATCT CTACAAACAT  
E--> 103  
W--> 104 TTAAAAAAAT TAGTCAGGTG AAGTGGTGCA TGGTGGTAGT CCCAGATATT  
E--> 105 TGGAAGGCTG AGGCAGGGAGG ATCGCTTGAG CCCAGGAATT TGAGGCTGCA  
E--> 106 GTGAGCTGTG ATCACACCAAC TGCACCTCCAG CCTCAGTGAC AGAGTGAGGC  
W--> 107 CCTGTCTCAA AAAAGAAAAG AAAAAGAAA AATAATGAGG GCTGTATGGA  
E--> 108 ATACGTTCAT TATTCAATTCA CTCACACT CACTCATTCA TTCATTCAATT  
W--> 109 CATTCAACAT GTCTTATTGC ATACCTTCTG TTTGCTCAGC TTGGTGCTTG  
E--> 110 AGAGTCCACT CCCTGTAGGT CGGGCAGCAG GCCGTAGAAG TCTGGCAGGG  
W--> 111 CCTGGCCCTG CTGTCGGAAG CTGTCCTGCG GGGCCAGGCC CTGTTGGTCA  
E--> 112 ACTCTTCCCA GCCGTGGGAG CCCCTGCAGC TGCATGTGGA TAAAGCCGTC  
W--> 113 AGTGGCCTTC GCAGCCTCAC CACTCTGCTT CGGGCTCTGG GAGCCCAGGT  
E--> 114 GAGTAGGAGC GGACACTTCT GCTTGCCTT TCTGTAAGAA GGGGAGAAGG  
W--> 115 GTCTTGCTAA GGAGTACAGG AACTGTCCGT ATTCCCTTCCC TTTCTGTGGC  
E--> 116 ACTGCAGCGA CCTCCTGTT TCTCCTTGGC AGAAGGAAGC CATCTCCCCT  
W--> 117 CCAGATGCGG CCTCAGCTGC TCCACTCCGA ACAATCACTG CTGACACTTT  
E--> 118 CCGCAAACTC TTCCGAGTCT ACTCCAATT CCTCCGGGGA AAGCTGAAGC  
W--> 119  
E--> 120  
W--> 121  
E--> 122  
W--> 123  
E--> 124  
W--> 125  
E--> 126  
W--> 127  
E--> 128  
W--> 129  
E--> 130  
W--> 131  
E--> 132  
W--> 133  
E--> 134  
W--> 135  
E--> 136  
W--> 137  
E--> 138

PAGE: 4

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/029,042ADATE: 10/28/1999  
TIME: 14:00:56

Input Set: I029042A.RAW

E--> 139 TGTACACAGG GGAGGCCTGC AGGACAGGGG ACAGATGA  
 W--> 140

---

141 <210> 3  
 142 <211> 1585  
 143 <212> DNA  
 144 <213> erythropoietin 3  
 E--> 145 <400> ATGGGGTGC ACCAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCCCTGCT  
 E--> 146 GTCGCTCCCT CTGGGCCTCC CAGTCCTGGG CGCCCCACCA CGCCTCATCT  
 W--> 147 GTGACAGCCG AGTCCTGGAG AGGTACCTCT TGGAGGCCAA GGAGGCCGAG  
 E--> 148  
 W--> 149 AATATCACGG TGAGACCCCT TCCCCAGCAC ATTCCACAGA ACTCACGCTC  
 E--> 150  
 W--> 151 AGGGCTTCAG GG(AACTCCT CCCAG(ATCC AGGAACCTGG CACTGGTTT  
 E--> 152  
 W--> 153 GGGGTGGAGT TGGGAAGCTA GACACTGCC CCCTACATAA GAATAAGTCT  
 E--> 154  
 W--> 155 GGTGGCCCA AACCATAACCT GGAAACTAGG CAAGGAGCAA AGCCAGCAGA  
 E--> 156  
 W--> 157 TCCTACGGCC TGTGGCCAG GGCCAA(AAC CTTCAGGGAC CCTTGACTCC  
 E--> 158  
 W--> 159 CCGGGCTGTT TGCATTTCAAG ACGGGCTGTG CTGAACACTG CAGCTTGAAT  
 E--> 160  
 W--> 161 GAAAATATCA CTGTCCCAGA CACCAAAGTT AATTCTATG CCTGGAAGAG  
 E--> 162  
 W--> 163 GATGGAGGTG AGTTCTTTT TTTTTTTTT TCCTTTCTTT TGGAGAATCT  
 E--> 164  
 W--> 165 CATTGCGAG CCTGATTTG GATGAAAGGG AGAATGATCG AGGGAAAGGT  
 E--> 166  
 W--> 167 AAAATGGAGC AGCAGAGATG AGGCTGCCTG GGCGCAGAGG CTCACGTCTA  
 E--> 168  
 W--> 169 TAATCCCAGG CTGAGATGGC CGAAATGGGA GAATTGCTTG AGCCCTGGAG  
 E--> 170  
 W--> 171 GTTCAGACCA ACCTAGGCAG CATACTGAGA TCCCCCATCT CTACAAACAT  
 E--> 172  
 W--> 173 TTAAAAAAAT TAGTCAGGTG AAGTGGTGCA TGGTGGTAGT CCCAGATATT  
 E--> 174  
 W--> 175 TGGATGGCTG AGGCGGGAGG ATCGCTTGAG CCCAGGAATT TGAGGCTGCA  
 E--> 176  
 W--> 177 GTGAGCTGTG ATCACACCAAC TGCACCTCCAG CCTCAGTGAC AGAATGAGGC  
 E--> 178  
 W--> 179 CCTGTCTCAA AAAAGAAAAG AAAAAGAAA AATAATGAGG GCTGTATGGA  
 E--> 180  
 W--> 181 ATACATTCA TATTCACTCA CTCACACT CACTCATCCA TTCATTCAATT  
 E--> 182  
 W--> 183 CATTCAACAA GTCTTATTGC ATACCTTCTG TTTGCTCAGC TTGGTGCTCG  
 E--> 184  
 W--> 185 GGGCTGCTGA GGGGCAGGAG GGAGAGGGTG ACATGGGTCA GCTGACTCCC  
 E--> 186  
 W--> 187

*Same*

PAGE: 5

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/029,042ADATE: 10/28/1999  
TIME: 14:00:56

Input Set: I029042A.RAW

E--> 188 AGAGTCCACT CCCTGTAGGT CGGGCAACAG GCCGTAGAAG TCTGGCAGGG  
 W--> 189 CCTGGCCCTG CTGTCGGAAG CTGTCCTGCG GGGCCAGGCC CTGTTGGTCA  
 E--> 190 ACTTTTCCCA GCCGTGGGAG CCCCTGCAGC TGCAATGTGGA TAAAGCCGTC  
 W--> 191 AGTGGCCTTC GCAGCCTCAC CACTCTGCTT CGGGCTCTGG GAGCCCAGGT  
 E--> 192 GAGTAGGAGC GGACACTTCT GCTTGCCCTT TCTGTAAGAA GGGGAGAAGG  
 W--> 193 GTCTTGCTAA GGAGTACAGG AACTGTCCGT ATTCCCTCCC TTTCTGTGGC  
 E--> 194 ACTGCAGCGA CCTCCTGTTT TCTCCTGGC AGAAGGAAGC CATCTCCCCT  
 W--> 195 CCAGATGCGG CCTCAGCTGC TCCACTCCGA ACAATCACTG CTGACACTTT  
 E--> 196 CCGCAAACTC TTCCGAGTCT ACTCCAATTT CCTCCGGGGA AAGCTGAAGC  
 W--> 197 TGTACACAGG GGAGGCCTGC AGGACAGGGG ACAGATGA  
 E--> 198  
 W--> 199  
 E--> 200  
 W--> 201  
 E--> 202  
 W--> 203  
 E--> 204  
 W--> 205  
 E--> 206  
 W--> 207

---

208 <210> 4  
 209 <211> 1585  
 210 <212> DNA  
 211 <213> erythropoietin 4  
 E--> 212 <400> ATGGGGGTGC AGGAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCCCTGCT  
 E--> 213 GTCGCTCCCT CTGGGCCTCC CAGTCCTGGG CGCCCCACCA CGCCTCATCT  
 W--> 214 GTGACAGCCG AGTCCTGGAG AGGTACCTCT TGGAGGCCAA GGAGGCCGAG  
 E--> 215  
 W--> 216 AATATCACGG TGAGACCCCT TCCCCAGCAC ATTCCACAGA ACTCACGCTC  
 E--> 217  
 W--> 218 AGGGCTTCAG GG AACTCCT CCCAG ATCC AGAACCTGG CACTTGGTTT  
 E--> 219  
 W--> 220 GGGGTGGAGT TGGGAAGCTA GACACTGCC CCCTACATAA GAATAAGTCT  
 E--> 221  
 W--> 222 GGTGGCCCCA AACCATACCT GGAAACTAGG CAAGGAGCAA AGCCAGCAGA  
 E--> 223  
 W--> 224 TCCTACGGCC TGTGGGCCAG GGCCAGGAGC CTTCAGGGAC CCTTGACTCC  
 E--> 225 CCGGGCTGTG TGCATTCAG AAGGGCTGTG CTGAACACTG CAGCTTGAAT  
 E--> 226  
 W--> 228 GAGAATATCA CTGTCCCAGA CACCAAAGTT AATTCTATG CCTGGAAGAG  
 E--> 229  
 W--> 230 GATGGAGGTG AGTTCTTTT TTTTTTTT TCCTTTCTTT TGGAGAATCT  
 E--> 231  
 W--> 232 CATTGCGAG CCTGATTTG GATGAAAGGG AGAGTGATCG AGGGAAAGGT  
 E--> 233  
 W--> 234 AAAATGGAGC AGCAGAGATG AGGCTGCCTG GGCGCAGAGG CTCACGTCTA  
 E--> 235  
 W--> 236

PAGE: 6

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/029,042ADATE: 10/28/1999  
TIME: 14:00:56

Input Set: I029042A.RAW

E--> 237 TAATCCCAGG CTGAGATGGC CGAGATGGGA GAATTGCTTG AGCCCTGGAG  
 W--> 238 GTTCAGACCA ACCTAGGCAG CATACTGAGA TCCCCCATCT CTACAAACAT  
 E--> 239 TTAAAAAAAT TAGTCAGGTG AAGTGGTGCA TGGTGGTAGT CCCAGATATT  
 W--> 240 TGGAAAGGCTG AGGCGGGAGG ATCGCTTGAG CCCAGGAATT TGAGGCTGCG  
 E--> 241 GTGAGCTGTG ATCACACCAC TGCACTCCAG CCTCAGTGAC AGAGTGAGGC  
 W--> 242 CCTGTCTCAA AAAAGAAAAG AAAAAGAAA AATAATGAGG GCTGTATGGA  
 E--> 243 ATACATTCA TATTCACTCA CTCACTCACT CACTCATTCA TTCATTCAATT  
 W--> 244 CATTCAACAA GTCTTATTGC ATACCTTCTG TTTGCTCAGC TTGGTGCTTG  
 E--> 245 GGGCTGCTGA GGGGCAGGAG GGTGAGGGTG ACATGGGTCA GCTGACTCCC  
 W--> 246 AGAGTCCACT CCCTGTTGGT CGGGCAGCAG GCCGTAGAAG TCTGGCAGGG  
 E--> 247 CCTGGCCCTG CTGTCGGAAG CTGTCCTGCG GGGCCAGGCC CTGTTGGTCA  
 W--> 248 ACTCTTCCA GCCGTGGAG CCCCTGCAGC TGCATGTGGA TAAAGCCGTC  
 E--> 249 AGTGGCCTTC GCAGCCTCAC CACTCTGCTT CGGGCTCTGG GAGCCCAGGT  
 W--> 250 GAGTAGGAGC GGACACTTCT GCTTGCCTT TCTGTAAGAA GGGGAGAAGG  
 E--> 251 GTCTTGCTAA GGAGTACAGG AACTGTCCGT ATTCCCTCCC TTTCTGTGGC  
 W--> 252 ACTGCAGCGA CCTCCTGTT TCTCCTGGC AGAAGGAAGC CATCTCCCCT  
 E--> 253 CCAGATGCGG CCTCAGCTGC TCCACTCCGA ACAATCACTG CTGACACTTT  
 W--> 254 CCGCAAACTC TTCCGAGTCT ACTCCAATTT CCTCCGGGAA AAGCTGAAGC  
 E--> 255 TGTACACAGG GGAGGCCTGC AGGACAGGGG ACAGATGA  
 W--> 256  
 E--> 257  
 W--> 258  
 E--> 259  
 W--> 260  
 E--> 261  
 W--> 262  
 E--> 263  
 W--> 264  
 E--> 265  
 W--> 266  
 E--> 267  
 W--> 268  
 E--> 269  
 W--> 270  
 E--> 271  
 W--> 272  
 E--> 273  
 W--> 274

275 &lt;210&gt; 5

276 &lt;211&gt; 1583

277 &lt;212&gt; DNA

278 &lt;213&gt; erythropoietin ) 5

*Danle*

E--> 279 <400> ATGGGGGTGC ACGAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCCCTGCT  
 E--> 280 GTCGCTCCCT CTGGGCCTCC CAGTCCTGGG CGCCCCACCA CGCCTCATCT  
 W--> 281 GTGACAGACG AGTCCTGGAG AGGTACCTCT TGGAGGCCAA GGAGGCCAG  
 E--> 282  
 W--> 283  
 E--> 284 AATATCACGG TGAGACCCCT TCCCCAGCAC ATTCCACAGA ACTCACGCTC  
 W--> 285

PAGE: 7

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/029,042ADATE: 10/28/1999  
TIME: 14:00:56

Input Set: I029042A.RAW

E--> 286 AGGGCTTCAG GG- AACTCCT CCCAG- ATCC AGGAACCTGG CACTTGGTTT  
 W--> 287 GGGGTGGAGT TGGGAAGCTA GACACTGCC CCCTACATAA GAATAAGTCT  
 E--> 288 GGTGGCCCCA AACCATAACCT GGAAACTAGG CAAGGAGCAA AGCCAGCAGA  
 W--> 289 TCCTACGGCC TGTGGCCCAG GGGCA- GAGC CTTCAGGGAC CCTTGACTCC  
 E--> 290 CCGGGCTGTG TGCATTTCA GCGGGCTGTG CTGAACACTG CAGCTTGAAT  
 W--> 291 GAGAATATCA CTGTCCCAGA CACCAAAGTT AATTCTATG CCTGGAAGAG  
 E--> 292 GATGGAGGTG AGTCCTTTT TTTTTTT- TCCTTCTTT TGGAGAATCT  
 W--> 293 CATTTGCGAG CCTGATTTG GATGAAAGGG AGAATGATCG AGGGAAAGGT  
 E--> 294 AAAATGGAGC AGCAGAGATG AGGCTGCCTG GGCGCAGAGG CTCACGTCTA  
 W--> 295 TAATCCCAGG CTGAGACGGC CGAGATGGGA GAATTGCTTG AGCCCTGGAG  
 E--> 296 GTTCAGACCA ACCTAGGCAG CATACTGAGA TCCCCCATCT CTACAAACAT  
 W--> 297 TTAAAAAAAT TAGTCAGGTG AAGTGGTGCA TGGTGGTAGT CCCAGATATT  
 E--> 298 TGGAAGGCTG AGGCAGGAGG ATCGCTTGAG CCCAGGAATT TGGGCTGCA  
 W--> 299 GTGAGCTGTG ATCACACCCAC TGCAATCCAG CCTCAGTGAC AGAGTGAGGC  
 E--> 300 CCTGTCTCAA AAACGAAAAG AAAAAGAAA ATAATGAGG GCTGTATGGA  
 W--> 301 ATACATTCA TATTCAATTCA CTCACTCACT CACTCATTCA TTCATTCAATT  
 E--> 302 CATTCAACAA GTCTTATTGC ATACCTTCTG TTTGCTCAGC TTGGTGTGTTG  
 W--> 303 GCCCTCTGA GGGGCAGGAG GGAGAGGGTG ACATGGGTCA GCTGACTCCC  
 E--> 304 AGAGTCCACT CCCTGTAGGT CGGGCAGCAG GCCGTAGAAG TCTGGCAGGG  
 W--> 305 CCTGGCCCTG CTGTGGAAT CTGTCTGCG GGGCCAGGCC CTGTTGGTCA  
 E--> 306 ACTCTTCCCA ACCGTGGGAG CCCCTGCAGC TGCAATGTGGA TAAAGCCGTC  
 W--> 307 AGTGGCCTTC GCAGCCTCAC CACTCTGCTT CGGGCTCTGG GAGCCCAGTT  
 E--> 308 GAGTAGGAGG GGACACTTCT GCTTGCCTT TGTGTAAGAA GGAGAGAAGG  
 W--> 309 GTCTTGCTAA GGAGTACAGG AACTGTCCGT ATTCCCTCCC TTTCTGTGGC  
 E--> 310 ACTGCAGCGA CCTCCTGTT TCTCCTTGGC AGAAGGAAGC CATCTCCCCT  
 W--> 311  
 E--> 312  
 W--> 313  
 E--> 314  
 W--> 315  
 E--> 316  
 W--> 317  
 E--> 318  
 W--> 319  
 E--> 320  
 W--> 321  
 E--> 322  
 W--> 323  
 E--> 324  
 W--> 325  
 E--> 326  
 W--> 327  
 E--> 328  
 W--> 329  
 E--> 330  
 W--> 331  
 E--> 332  
 W--> 333  
 E--> 334  
 W--> 335

PAGE: 8

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/029,042A

DATE: 10/28/1999  
TIME: 14:00:56

Input Set: I029042A.RAW

E--> 336	CCAGATGCGG CCTCAGCTGC TCCACTCCGA ACAATCACTG CTGATACTTT
W--> 337	CCGCAAACTC TTCCGAGTCT ACTCCAATT CCTCCGGGAA AAGCTGAAGC
E--> 338	
W--> 339	
E--> 340	TGTACACAGG GGAGGCCTGC AGGACAGGGG ACAGATGA
W--> 341	

---

342 <210> 6	<i>Same</i>
343 <211> 1586	
344 <212> DNA	
345 <213> erythropoietin → 6	
E--> 346 <400> ATGGGGGTGC ACGAATGTCC TGCCTGGCTG TGGCTTCTCC TGTCCCTGCT	
E--> 347 GTCGCTCCCT CTGGGCCTCC CAGTCCTGGG CGCCCCACCA CGCCTCATCT	
W--> 348 GTGACAGCCG AGTCCTGGAG AGGTACCTCT TGGAGGCCAA GGAGGCCGAG	
E--> 349	
W--> 350 AATATCACGG TGAGACCCCT TCCCCAGCAC ATTCCACAGA ACTCACGCTC	
E--> 351	
W--> 352 AGGGCTTCAG GGGAACTCCT CCCAGGATCC AGGAACCTGG CACTTGGTTT	
E--> 353	
W--> 354 GGGGTGGAGT TGGGAAGCTA GACACTGCC CCCTACATAA GAATAAGTCT	
E--> 355	
W--> 356 GGTGGCCCCA AACCATACCT GGAAACTAGG CAAGGAGCAA AGCCAGCAGA	
E--> 357	
W--> 358 TCCTACGGCC TGTGGCCAG GGCA <del>G</del> AGC CTTCAGGGAC CCTTGACTCC	
E--> 359	
W--> 360 CCGGGCTGTG TGCATTCCAG ACGGGCTGTG CTGAACACTG CAGCTTGAAT	
E--> 361	
W--> 362 GAGAATATCA CTGTCCCAGA CACCAAAGTT AATTCTATG CCTGGAAGAG	
E--> 363	
W--> 364 GATGGAGGTG AGTTCTTTT TTTTTTTTT TCCTTTCTTT TGGAGAATCT	
E--> 365	
W--> 366 CATTGCGAG CCTGATTTGG GATGAAAGGG AGAATGATCG AGGGAAAGGT	
E--> 367	
W--> 368 AAAATGGAGC AGCAGAGATG AGGCTGCCTG GGCGCAGAGG CTCCAGTCTA	
E--> 369	
W--> 370 TAATCCCAGG CTGAGATGGC CGAGATGGGA GAATTGCTTG AGCCCTGGAG	
E--> 371	
W--> 372 GTTCAGACCA ACCTAGGCAG CCTAGTGAGA TCCCCCATCT CTACAAACAT	
E--> 373	
W--> 374 TTAAAAAAAT TAGTCAGGTG AAGTGGTGCA TGGTGGTAGT CCCAGATATT	
E--> 375	
W--> 376 TGGAAGGCTG AGGCGGGAGG ATCGCTTGAG CCCAGGAATT TGAGGCTGCA	
E--> 377	
W--> 378 GTGAGCTGTG ATCACACCAC TGCACCTCCAG CCTCAGTGAC AGAGTGAGGC	
E--> 379	
W--> 380 CCTGTCTCAA AAAAGAAAAG AAAAAGAAA AATTATGAGG GCTGTATGGA	
E--> 381	
W--> 382 ATACATTCA TATTCACTCA CTCACACT CACTCATTCA TTCATTCA	
E--> 383	
W--> 384	

PAGE: 9  
**RAW SEQUENCE LISTING**  
 PATENT APPLICATION US/09/029,042A

DATE: 10/28/1999  
 TIME: 14:00:56

Input Set: I029042A.RAW

E--> 385	CATTCAACAA GTCTTATTGC ATACCTTCTG TTTGCTCAGC TTGGTGCTTG
W--> 386	GGGCTGCTGA GGGCAGGAG GGAGAGGGTG ACATGGGITC ACTGACTCCC
E--> 387	AGAGTCCACT CCCTGTAGGT CGGGCAGCAG GCCGTAGAAG TCTGGCAGGG
W--> 388	CCTGGCCCTG CTGTCGGAAG CTGTCCTGCG GGGCCAGGCC CTGTTGGTCA
E--> 389	ACTCTTCCA GCCGTGGAG CCCCTGCAGC TGCAATGTGGA TAAAGCCGTC
W--> 390	AGTGGCCTTC GCAGCCTCAC CACTCTGCTT CGGGCTCTGG GAGCCCAGGT
E--> 391	GAGTAGGAGC GGACACTTCT GCTTGCCCTT TCTGTAAGAA GGGGAGAAGG
W--> 392	GTCTTGCTAA GGAGTACAGG ATCTGTCCGT ATTCCCTTCCC TTTCTGTGGC
E--> 393	ACTGCAGCGA CCACCTGTTT TCTCCTTGGC AGAAGGAAGC CATCTCCCCT
W--> 394	CCAGATGCGG CCTCAGCTGC TCCACTCCGA ACAATCACTG CTGACACTTT
E--> 395	CCGCAAACTC TTCCGAGTCT ACTCCAATTT CCTCCGGGGA GAGCTGAAGC
W--> 396	TGTACACAGG GGAGGCCTGC AGGACAGGGG ACGGATGA
E--> 397	
W--> 398	
E--> 399	
W--> 400	
E--> 401	
W--> 402	
E--> 403	
W--> 404	
E--> 405	
W--> 406	
E--> 407	
W--> 408	

E--> 409	<210> 7	<i>do not show amino acids next to &lt;400&gt;</i>										
E--> 410	<211> 193											
E--> 411	<212> PRT											
E--> 412	<213> erythropoietin gene	7										
E--> 413	<400> Met Gly Val His Glu Cys Pro Ala Trp Leu Trp Leu Leu Ser Leu	15	<i>per 1.822 of sequence length, number the amino acids under every 5 amino acids.</i>									
E--> 414	Leu Ser Leu Pro Leu Gly Leu Pro Val Leu Gly Ala Pro Pro Arg Leu	25										
E--> 415	Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu											
E--> 416												
E--> 417												
E--> 418												
E--> 419												
E--> 420												
E--> 421												
E--> 422												
E--> 423												
E--> 424												
E--> 425												
E--> 426												
E--> 427												
E--> 428												
E--> 429												
E--> 430												
E--> 431												
E--> 432												
E--> 433												

*why are just 2 amino acids shown on the line?*

*Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile*

*Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu*

*Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Ser*

*Glu Ala*

*Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu*

*Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu*

*Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro*

*per 1.822*

*of*

*sequence  
length,*

*number*

*the amino  
acids  
under*

*every 5  
amino acids.*

*Do NOT use  
TAB codes*

*between  
amino acid nos.*

*Use space  
characters.*

PAGE: 10

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION US/09/029,042A**

DATE: 10/28/1999  
 TIME: 14:00:56

Input Set: I029042A.RAW

434                  Pro Asp

E--> 435                  Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg

E--> 436                  Lys Leu Phe Arg Val Val Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu

E--> 437                  Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg

E--> 438

E--> 439

E--> 440

E--> 441

---

442 <210> 8                  same error

E--> 443 <211> 193

E--> 444 <212> PRT

E--> 445 <213> erythropoietin gene                  8

E--> 446 <400> Met Gly Val His Glu Cys Pro Ala Trp Leu Trp Leu Leu Ser Leu

E--> 447                  Leu Ser Leu Pro Leu Gly Leu Pro Val Leu Gly Ala Pro Pro Arg Leu

E--> 448                  Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu

E--> 449

E--> 450                  Ala Glu

E--> 451

E--> 452                  Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile

E--> 453

E--> 454                  Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile

E--> 455                  Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu

E--> 456                  Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser

E--> 457

E--> 458                  Glu Ala

E--> 459

E--> 460                  Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu

E--> 461

E--> 462                  Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu

E--> 463

E--> 464                  Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro

E--> 465

E--> 466                  Pro Asp

E--> 467

E--> 468                  Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg

E--> 469

E--> 470                  Lys Leu Phe Arg Val Val Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu

E--> 471

E--> 472                  Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg

E--> 473

E--> 474

---

E--> 475 <210> 9                  number  
 E--> 476 <211> 193                  and  
 E--> 477 <212> PRT                  and  
 E--> 478 <213> erythropoietin gene                  9

E--> 479 <400> Met Gly Val His Glu Cys Pro Ala Trp Leu Trp Leu Leu Ser Leu

E--> 480                  Leu Ser Leu Pro Leu Gly Leu Pro Val Leu Gly Ala Pro Pro Arg Leu

E--> 481

---

number  
 the  
 and  
 and  
 every 5  
 and  
 and  
 Do not  
 use TAB  
 codes.

number  
 and  
 and

PAGE: 11

**RAW SEQUENCE LISTING**  
**PATENT APPLICATION US/09/029,042A**

DATE: 10/28/1999  
 TIME: 14:00:56

Input Set: I029042A.RAW

482 Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu  
 E--> 483  
 484 **Ala Glu**  
 E--> 485 Asn Ile Thr Lys Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile  
 486  
 E--> 487 Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu  
 488  
 E--> 489 Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser  
 490  
 E--> 491 Glu Ala  
 492  
 E--> 493 Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu  
 494  
 E--> 495 Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu  
 496  
 E--> 497 Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro  
 498  
 E--> 499 **Pro Asp**  
 500  
 E--> 501 Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg  
 502  
 E--> 503 Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu  
 504  
 E--> 505 Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg  
 506

---

507 <210> 10  
 E--> 508 <211> 193  
 509 <212> PRT  
 510 <213> erythropoietin gene → /O  
 E--> 511 <400> Met Gly Val His Glu Cys Pro Ala Trp Leu Trp Leu Leu Leu Ser Leu  
 512 Leu Ser Leu Pro Leu Gly Leu Pro Val Leu Gly Ala Pro Pro Arg Leu  
 513 Ile Cys Asp Arg Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu  
 514  
 E--> 515 Ala Glu  
 516  
 E--> 517 Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile  
 518  
 E--> 519 Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu  
 520  
 E--> 521 Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser  
 522  
 E--> 523 **Glu Ser**  
 524  
 E--> 525 Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu  
 526  
 E--> 527 Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu  
 528  
 E--> 529 Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro  
 530

PAGE: 12

RAW SEQUENCE LISTING  
PATENT APPLICATION US/09/029,042ADATE: 10/28/1999  
TIME: 14:00:56

Input Set: I029042A.RAW

E--> 531                          Pro Asp

E--> 532

E--> 533                          Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg

E--> 534

E--> 535                          Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Lys Leu Lys Leu

E--> 536

E--> 537                          Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Arg

E--> 538

E--> 539

---

E--> 540 <210> 11

E--> 541 <211> 193

E--> 542 <212> PRT

E--> 543 <213> erythropoietin gene >/ /

E--> 544 <400> Met Gly Val His Glu Cys Pro Ala Trp Leu Trp Leu Leu Ser Leu

E--> 545                          Leu Ser Leu Pro Leu Gly Leu Pro Val Leu Gly Ala Pro Pro Arg Leu

E--> 546                          Ile Cys Asp Ser Arg Val Leu Glu Arg Tyr Leu Leu Glu Ala Lys Glu

E--> 547

E--> 548                          Ala Glu

E--> 549

E--> 550                          Asn Ile Thr Thr Gly Cys Ala Glu His Cys Ser Leu Asn Glu Asn Ile

E--> 551

E--> 552                          Thr Val Pro Asp Thr Lys Val Asn Phe Tyr Ala Trp Lys Arg Met Glu

E--> 553

E--> 554                          Val Gly Gln Gln Ala Val Glu Val Trp Gln Gly Leu Ala Leu Leu Ser

E--> 555

E--> 556                          557                          Glu Ala

E--> 558                          Val Leu Arg Gly Gln Ala Leu Leu Val Asn Ser Ser Gln Pro Trp Glu

E--> 559

E--> 560                          Pro Leu Gln Leu His Val Asp Lys Ala Val Ser Gly Leu Arg Ser Leu

E--> 561

E--> 562                          Thr Thr Leu Leu Arg Ala Leu Gly Ala Gln Lys Glu Ala Ile Ser Pro

E--> 563

E--> 564                          565                          Pro Asp

E--> 566                          Ala Ala Ser Ala Ala Pro Leu Arg Thr Ile Thr Ala Asp Thr Phe Arg

E--> 567

E--> 568                          Lys Leu Phe Arg Val Tyr Ser Asn Phe Leu Arg Gly Glu Leu Lys Leu

E--> 569

E--> 570                          Tyr Thr Gly Glu Ala Cys Arg Thr Gly Asp Gly

E--> 571

E--> 572                          1'

(FU)  
delete at end of file

Please review the

Sequence Listing to ensure that a corresponding explanation is presented in the <220> to  
<223> fields of each sequence which presents at least one n or Xaa.

Appendix A To Subpart C to Part 1—Sample Sequence Listing

<110> Smith, John

Smith, Jane

<120> Example of a Sequence Listing

<130> 01-00001

<140> US 08/999,999

<141> 1998-02-28

<150> EP 91000000

<151> 1997-12-31

Please consult.

<160> 2

<170> PatentIn ver. 2.0

<210> 1

<211> 403

<212> DNA

<213> Paramecium aurelia

<220>

<221> CDS

<222> 341..394

<300>

<301> Doe, Richard

<302> Isolation and Characterization of a Gene Encoding a  
Protease from Paramecium sp.

<303> Journal of Fictional Genes

<304> 1

<305> 4

<306> 1 - 7

<307> 1988-06-20

<400> 1

ctactctact ctactctcat ctactatctt ctttgatct ctgagtctgc ctgagtgta 60

ctcttgatc ctggagatct ctcctctcac atgtgatcgt cgagactgac cgatagatcg 120

ctgactgact ctgagatagt cgagccgta cgagaccgt cgagggtgac agagagtggg 180

cgcgtgcgcg cagagcgccg cgccggtgcg cgcgcgagtg cgccgtgggc cgcgcgagg 240

cttcgcggc agcggcggcg ctteccggcg cgcccccgtc cgccctaga cctgagaggt 300

cttctttcc ctcctttca ctagagaggt ctatataac atg gtt tca atg ttc 355

Met Val Ser Met Phe

agc ttg tct ttc aaa tgg cct gga ttt tgt ttg ttt gtt tgtttgctc 403

Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu Phe Val

- 10

15

<210> 2

<211> 18

<212> PRT

<213> Paramecium aurelia

<400> 2

Met Val Ser Met Phe Ser Leu Ser Phe Lys Trp Pro Gly Phe Cys Leu

1

5

10

15

Phe Val

ed: May 22, 1998.

A. Lehman,  
ant Secretary of Commerce and  
issioner of Patents and Trademarks.  
oc. 98-14194 Filed 5-29-98; 8:45 am]  
CODE 3510-16-C

identifiers and their accompanying information as shown in the following table. The numeric identifier shall be used only in the "Sequence Listing." The order and presentation of the items of information in the "Sequence Listing" shall conform to the arrangement given below. Each item of information shall begin on a new line and shall begin with the numeric identifier enclosed in angle brackets as shown. The submission of those items of information designated with an "M" is mandatory. The submission of those items of information designated with an "O" is optional. Numeric identifiers <110> through <170> shall only be set forth at the beginning of the "Sequence Listing." The following table illustrates the numeric identifiers.

Numeric Identifier	Definition	Comments and Format	Mandatory (M) or Optional (O)
<110>	Applicant	'Preferably max. of 10 names; one name per line; preferable format: Surname, Other Names and/or Initials	M
<120>	Title of Invention		M
<130>	File Reference	Personal file reference	M when filed prior to assignment of appl. number
<140>	Current Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if available
<141>	Current Filing Date	Specify as: yyyy-mm-dd	M, if available
<150>	Prior Application Number	Specify as: US 07/999,999 or PCT/US96/99999	M, if applicable include priority documents under 35 USC 119 and 120
<151>	Prior Application Filing Date	Specify as: yyyy-mm-dd	M, if applicable
<160>	Number of SEQ ID NOS	Count includes total number of SEQ ID NOS	M
<170>	Software	Name of software used to create the Sequence Listing	O
<210>	SEQ ID NO:#:	Response shall be an integer representing the SEQ ID NO shown	M
<211>	Length	Respond with an integer expressing the number of bases or amino acid residues	M

<212>	Type	Whether presented sequence molecule is DNA, RNA, or PRT (protein). If a nucleotide sequence contains both DNA and RNA fragments, the type shall be "DNA." In addition, the combined DNA/RNA molecule shall be further described in the <220> to <223> feature section.	M
<213>	Organism	Scientific name, i.e. Genus/species, Unknown or Artificial Sequence. In addition, the "Unknown" or "Artificial Sequence" organisms shall be further described in the <220> to <223> feature section.	M
<220>	Feature	Leave blank after <220>. <221-223> provide for a description of points of biological significance in the sequence.	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<221>	Name/Key	Provide appropriate identifier for feature, preferably from WIPO Standard ST.25 (1998), Appendix 2, Tables 5 and 6	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence
<222>	Location	Specify location within sequence; where appropriate state number of first and last bases/amino acids	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified

		in feature	base was used in a sequence
<223>	Other Information	Other relevant information; four lines maximum	M, under the following conditions: if "n," "Xaa," or a modified or unusual L-amino acid or modified base was used in a sequence; if ORGANISM is "Artificial Sequence" or "Unknown"; if molecule is combined DNA/RNA.
<300>	Publication Information	Leave blank after <300>	O
<301>	Authors	Preferably max of ten named authors of publication; specify one name per line; preferable format: Surname, Other Names and/or Initials	O
<302>	Title		O
<303>	Journal		O
<304>	Volume		O
<305>	Issue		O
<306>	Pages		O
<307>	Date	Journal date on which data published; specify as yyyy-mm-dd, MMM-yyyy or Season-yyyy	O
<308>	Database Accession Number	Accession number assigned by database including database name	O
<309>	Database Entry Date	Date of entry in database; specify as yyyy-mm-dd or MMM-yyyy	O
<310>	Patent Document Number	Document number; for patent-type citations only. Specify as, for example, US 07/999,999	O

<311>	Patent Filing Date	Document filing date, for patent-type citations only; specify as yyyy-mm-dd	O
<312>	Publication Date	Document publication date, for patent-type citations only; specify as yyyy-mm-dd	O
<313>	Relevant Residues	FROM (position) TO (position)	O
<400>	Sequence	SEQ ID NO should follow the numeric identifier and should appear on the line preceding the actual sequence	M

5. Section 1.824 is revised to read as follows:

1.824 Form and format for nucleotide and/or amino acid sequence submissions in computer readable form.

(a) The computer readable form required by 1.821(e) shall meet the following specifications:

(1) The computer readable form shall contain a single "Sequence Listing" as either a diskette, series of diskettes, or other permissible media outlined in paragraph (c) of this section.

(2) The "Sequence Listing" in paragraph (a) (1) of this section shall be submitted in American Standard Code for Information Interchange (ASCII) text. No other formats shall be allowed.

(3) The computer readable form may be created by any means, such as word processors, nucleotide/amino acid sequence editors or other custom computer programs; however, it shall conform to all specifications detailed in this section.

(4) File compression is acceptable when using diskette media, so long as the compressed file is in a self-extracting format that will decompress on one of the systems described in paragraph (b) of this section.

(5) Page numbering shall not appear within the computer readable form version of the "Sequence Listing" file.

(6) All computer readable forms shall have a label permanently affixed thereto on which has been hand-printed or typed: the name of the applicant, the title of the invention, the date on which the data were recorded on the computer readable form, the operating system used, a reference number, and an application serial number and filing date, if known.

(b) Computer readable form submissions must meet these format requirements:

(1) Computer: IBM PC/XT/AT, or compatibles, or Apple Macintosh;

(2) Operating System: MS-DOS, Unix or Macintosh;